

## RECOMMENDATIONS

Five recommendations are provided here from the Matching Group, Administrative Records Subcommittee, Federal Committee on Statistical Methodology. The first recommendation calls for the establishment of a continuing interagency working group on record linkage systems and techniques: such a working group would be expected to play a significant role in implementing recommendations 2 through 5. The second recommendation calls for careful monitoring of external developments that might affect the prospects for undertaking record linkages for statistical purposes. Recommendations 3, 4 and 5 identify specific aspects of record linkage systems and techniques that deserve special emphasis in future research, development and evaluation activities. The five recommendations are:

1. Documentation should be improved and information on record linkage systems and techniques should be shared.

It is recommended that the Matching Group of the Administrative Records Subcommittee be reconstituted as a Technical Working Group on Record Linkage Systems and Techniques, continuing to function under the auspices of the Federal Committee on Statistical Methodology. The main goal of the Working Group would be to promote the effective use of record linkage techniques for statistical purposes by encouraging the documentation of individual record linkage systems and techniques and the sharing of relevant technical information. A primary activity would be sponsorship and organization of workshops and meetings of professional societies to discuss relevant new developments and research, and to disseminate information on existing systems and techniques. In addition, the reconstituted working group would contribute, in appropriate ways, to the implementation of recommendations 2 through 5 below.

2. Changes in the external environment for record linkages should be monitored.

Statistical users of record linkage techniques should track external developments that may influence their ability to perform record linkages. Such developments include changes in laws, regulations and policies affecting access to records and changes in the content of data files used in record linkages. Examples of the latter would include increased use of four-digit ZIP code add-ons ("ZIP + 4") and steps taken to promote the use of unique addresses in rural areas. In so far as possible, statistical users of record linkage techniques, working through the reconstituted Working Group (see recommendation 1), should attempt to influence the

course of these developments in ways that will facilitate statistical applications. For example, the Working Group might try to promote the development of standards for reporting names and addresses of both businesses and individuals.

3. Comparative evaluation studies of record linkage systems should be undertaken.

Several agencies of the United States and Canadian governments have invested substantial resources in the development of automated record linkage systems for use in a variety of statistical programs. For many new applications, use of an existing system is likely to be more cost-effective than development of a new one. To aid potential users of record linkage systems, it is recommended that resources be sought for comparative evaluations of existing systems and some of their components, such as name and address standardizers and blocking rules. The evaluation design should recognize that record linkage systems vary in their objectives, especially with respect to the kinds of units for which records are to be matched: persons or businesses. A much-needed first step is the development of a detailed evaluation plan that specifies the measures of quality and cost to be used in the evaluation and the nature of the files to be matched. Such evaluations may require data sets for which true match status is known. One possibility would be to create such data sets by simulation.

4. Research and development aimed at the improvement of record linkage systems and techniques should give priority to selected aspects.

Recognizing that resources for the development of improved record linkage systems are limited, it is recommended that priority be given to the following aspects: (1) systems for linking business records, (2) name and address standardizers, (3) string comparators, (4) the choice of blocking strategies, (5) the development of "learning" systems, and (6) the role of manual intervention.

5. Errors associated with record linkages and their effects on analyses should be measured.

It is recommended that more research be carried out on the error characteristics of record linkage systems and on the effects of errors on analyses performed with the linked data sets. To enhance the value of such research, consensus is desirable on standard measures of record linkage errors and on methods of measuring them. Promising error measurement methods include

multiple matching techniques and direct contacts with samples of linked pairs to determine their true match status.

By design, the principal focus of the Workshop discussions and followup comments by participants was on methodological aspects of record linkages for statistical purposes. Legal and ethical considerations in such linkages were not part of the main agenda.

Nevertheless, the Matching Group of the Administrative Records Subcommittee recognizes that legal and ethical considerations must be weighed carefully by any organization that links

records from different sources and that public perceptions of the appropriateness of various kinds of record linkages are also of critical importance. More research in these areas would also be desirable, addressing, in particular: (1) public understanding of and attitudes toward linkages performed for statistical and other purposes; (2) survey respondents' comprehension of informed consent statements currently being used, especially when such statements cover linkages of survey data and administrative records; and (3) the effects on survey response of varying the amount and kinds of information included in informed consent statements to respondents.

Members of the  
MATCHING GROUP: ADMINISTRATIVE RECORDS SUBCOMMITTEE  
(as of May 1985)

Mary Bentz  
Internal Revenue Service

David Cordray  
General Accounting Office

Jane Crane  
National Center for Education Statistics

Maria Elena Gonzalez  
Office of Management and Budget

Nick Greenia  
Internal Revenue Service

David Hirschberg  
Small Business Administration

Howard Hogan  
Bureau of the Census

Thomas B. Jabine  
Consultant, Committee on National Statistics

Matt Jaro  
Bureau of the Census

Patrick Kelley  
Bureau of the Census

Beth Kilss  
Internal Revenue Service

Nancy Kirkendall  
Energy Information Administration

Douglas Kleweno  
Department of Agriculture

Thomas Reilly  
Bureau of the Census

Douglas Sater  
Bureau of the Census

Fritz Scheuren (Chair)  
Internal Revenue Service

Wray Smith  
Mathematica Policy Research

Wendel Thompson  
Department of Energy

Carol Utter  
Bureau of Labor Statistics

William Winkler  
Energy Information Administration